Journal of Geographic Information System, 2025, 17(6), 400-421
https://www.scirp.org/journal/jgis
ISSN Online: 2151-1969
ISSN Print: 2151-1950

Machine Learning-Assisted Assessment of PPK-Based UAV Mapping Accuracy in Complex Terrain: A Case Study of Umiam Reservoir, Meghalaya, India

Victor Saikhom[1*], Dibyajyoti Chutia[1], Sanjay Pandit[1], Shiv P. Aggrawal[1], Manoranjan Kalita[2]

[1]North Eastern Space Applications Centre, Department of Space, Government of India, Umiam, India
[2]Department of Civil Engineering, Assam Don Bosco University, Guwahati, India
Email: *victor.saikhom@nesac.gov.in

How to cite this paper: Saikhom, V., Chutia, D., Pandit, S., Aggrawal, S.P. and Kalita, M. (2025) Machine Learning-Assisted Assessment of PPK-Based UAV Mapping Accuracy in Complex Terrain: A Case Study of Umiam Reservoir, Meghalaya, India. *Journal of Geographic Information System*, **17**, 400-421.
https://doi.org/10.4236/jgis.2025.176019

Received: November 14, 2025
Accepted: December 26, 2025
Published: December 29, 2025

Abstract

Unmanned Aerial Vehicle (UAV) Photogrammetry has become an important tool for high-resolution mapping, with workflows increasingly incorporating onboard GNSS solutions such as Post-Processed Kinematic (PPK) to reduce dependence on Ground Control Points (GCPs). However, the performance of PPK in complex terrain remains uncertain. This study evaluates the positional accuracy of UAV-derived orthomosaics generated using GCP-based and PPK-based workflows over the hilly terrain of the Umiam Reservoir, Meghalaya, India, and applies machine learning to analyze terrain-induced error patterns. Accuracy assessment against DGPS-surveyed ground truth showed that the GCP-based workflow achieved sub-decimeter accuracy ($RMSE_X$ = 0.043 m, $RMSE_Y$ = 0.031 m, $RMSE_Z$ = 0.072 m; CE90 = 0.080 m; LE90 = 0.119 m). In contrast, the PPK-based workflow displayed larger deviations ($RMSE_X$ = 1.025 m, $RMSE_Y$ = 1.240 m, $RMSE_Z$ = 2.980 m; CE90 = 2.441 m; LE90 = 4.902 m), with particularly high vertical errors. Machine Learning (ML) provided additional insights: regression analysis revealed that slope explained ~60% of the variance in vertical error (R^2 = 0.605), while clustering identified three error regimes, *i.e.*, systematic vertical underestimation, systematic vertical overestimation, and horizontal drift. The results highlight a trade-off between field efficiency and accuracy. While PPK reduces GCP requirements, its vertical reliability declines sharply in rugged terrain. The novelty of this work lies in using ML-assisted diagnostics to characterize terrain-induced error regimes, offering insights beyond conventional RMSE-based evaluations. This study concludes that hybrid approaches, combining PPK with sparse GCPs or terrain-aware ML corrections, can balance efficiency and reliability. The findings have

practical significance for UAV mapping in the hilly landscapes of Northeast India and other complex terrains worldwide.

Keywords

UAV Photogrammetry, Post-Processed Kinematic (PPK), Ground Control Points (GCP), Geolocation Accuracy, Terrain Slope, Complex Terrain, Digital Surface Model (DSM)

1. Introduction

UAV-based Photogrammetry has rapidly emerged as one of the most significant innovations in geospatial sciences, enabling high-resolution data acquisition at user-defined temporal and spatial scales. The ability to flexibly deploy UAVs in remote or inaccessible regions has transformed surveying, monitoring, and mapping practices across multiple domains such as agriculture, forestry, hydrology, disaster management, and infrastructure development. A key strength of UAV systems lies in their capacity to capture ultra-high-resolution imagery with centimeter-level Ground Sampling Distances (GSDs), which are often not achievable with conventional airborne or satellite remote sensing platforms [1] [2].

In recent years, the integration of precise positioning systems such as Real-Time Kinematic (RTK) and Post-Processed Kinematic (PPK) Global Navigation Satellite Systems (GNSSs) has further enhanced the potential of UAV Photogrammetry. These techniques make it possible to achieve centimeter-level positional accuracy in orthomosaics and Digital Elevation Models (DEMs) with reduced reliance on extensive networks of GCPs [3] [4]. This represents a paradigm shift from conventional workflows that traditionally depended on multiple GCPs for accurate georeferencing. While GCPs provide reliable control, their collection is laborious, time-consuming, and often infeasible in hilly or forested terrain, particularly in regions like Northeast India, where dense vegetation and steep slopes impede field-based surveying [5] [6].

The PPK approach, in particular, has become a promising alternative by post-processing GNSS data against a fixed reference base station, thereby improving positional accuracy without requiring a dense GCP network [3] [4]. High-end UAV platforms such as the Trinity F90+, equipped with Vertical Take-Off and Landing (VTOL) capabilities and advanced full-frame sensors like the Sony RX1R II, allow long endurance flights and accurate geotagging of imagery in topographically challenging areas [7]. However, while PPK has been shown to deliver comparable accuracies to GCP-based workflows in flat or moderately undulating areas, its performance in complex, mountainous terrain remains less understood.

Topographic complexity introduces a unique set of challenges for UAV Photogrammetry. Steep slopes, varying aspects, and dense vegetation can reduce image overlap, complicate tie-point matching, and increase occlusions in Structure-from-Motion (SfM) processing [8]. Relief displacement is more severe in hilly ter-

rain, while GNSS signal obstruction and multipath effects are also more pronounced, which may result in degraded vertical accuracy. Consequently, while PPK workflows are operationally efficient, their robustness under such conditions requires rigorous evaluation. This gap is particularly evident in Meghalaya, India, where steep slopes and dense vegetation complicate UAV operations and ground surveying.

Beyond positional accuracy assessment through conventional statistical measures, recent advances in Artificial Intelligence (AI) and Machine Learning (ML) provide powerful tools to explore and interpret error behavior in UAV-derived datasets. Traditionally, UAV positional accuracy has been evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and confidence-level statistics such as CE90 and LE90 [9]. While effective, these methods may fail to capture non-linear interactions between terrain variables and positional errors. ML techniques, ranging from regression models to clustering algorithms, offer a complementary perspective by modeling hidden patterns in error distributions, identifying systematic biases, and predicting error magnitudes based on topographic and environmental conditions [10]-[12]. Even basic models, such as linear regression and K-means clustering, can provide valuable insights into the influence of slope, elevation, or land cover on UAV positional errors.

Globally, studies have begun exploring UAV-PPK performance and the influence of GCP placement strategies on photogrammetric accuracy. For example, Zhao *et al.* [4] investigated the impacts of GCP distribution in Greenland glaciers, while Seo *et al.* [13] examined optimal GCP placement strategies for UAV-based orthophoto production. Similarly, Martínez-Carricondo *et al.* [3] demonstrated the potential of PPK/RTK solutions across multiple GNSS base stations. Yet, there is a noticeable scarcity of research focusing on highly dissected, subtropical hilly landscapes, where UAV surveys are both most needed and most difficult to execute. In India, particularly the North Eastern Region (NER), UAV applications have mostly been confined to agriculture, forestry, and disaster monitoring [14]. Comprehensive positional accuracy studies that integrate PPK workflows with ML-based error analysis remain rare, leaving a significant gap in operational knowledge.

The Umiam Reservoir, located in Meghalaya's Ri-Bhoi district, offers an ideal setting to investigate these issues. The reservoir is a critical multipurpose water body, supporting hydroelectric generation, irrigation, and tourism, and is surrounded by steep hills, dense vegetation, and fluctuating water levels [13] [14]. These conditions pose significant challenges for conventional surveying, highlighting the practical importance of testing PPK-based UAV workflows. A systematic accuracy assessment in such terrain not only addresses a scientific gap but also has practical implications for watershed management, siltation monitoring, and infrastructure planning in the NER.

The present study contributes to this growing body of literature by: 1) comparing positional accuracy between PPK-based and GCP-based UAV orthomosaics

over a complex terrain landscape, 2) quantifying errors using standard RMSE, CE90, and LE90 metrics, and 3) integrating ML methods to analyze the relationship between terrain characteristics and error patterns. The novelty of this research lies in combining a state-of-the-art UAV platform with ML-assisted error analysis in a topographically challenging environment where limited prior work exists. By exploring both the operational benefits and limitations of PPK workflows, this study provides evidence-based recommendations for UAV mapping practices in hilly regions, with broader implications for geospatial applications in similar terrains worldwide.

2. Study Area

The study was conducted in and around the Umiam Reservoir, located in the Ri-Bhoi district of Meghalaya, India. This reservoir, also known as Barapani, is an important multipurpose water body used for hydroelectric power generation, irrigation, and tourism. The region is characterized by undulating terrain, steep slopes, dense vegetation, and significant seasonal variation in water levels, making it a representative site for evaluating UAV-based geospatial workflows in complex topographic conditions.

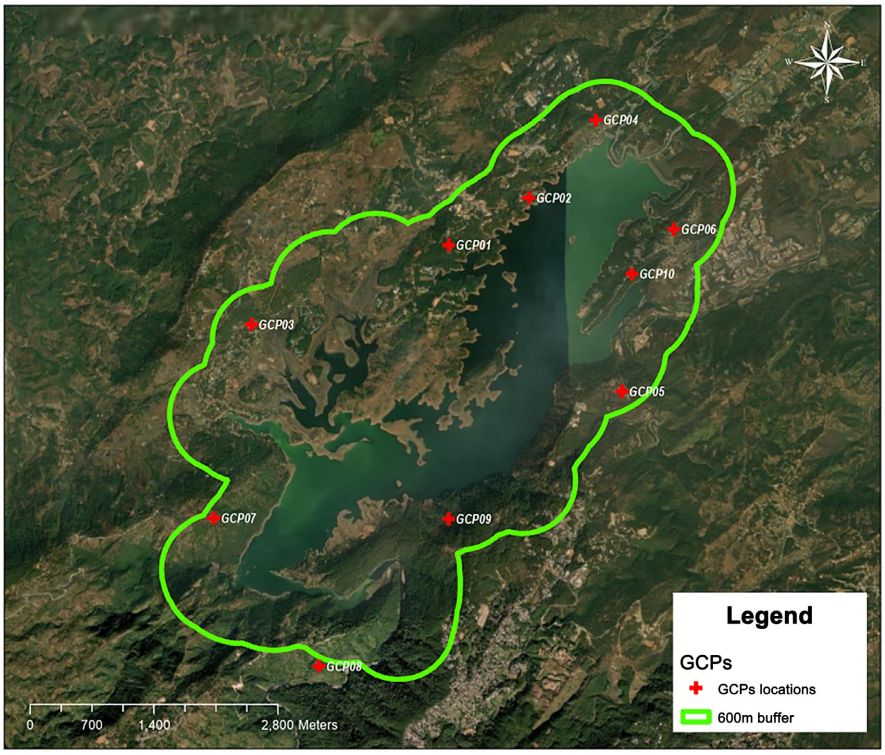


Figure 1. Study area map with 600 m buffer around Umiam Reservoir.

For the purpose of this study, a 600-meter buffer zone around the reservoir boundary was defined as the area of interest. This buffer covers a range of landscape features, including water surfaces, shoreline, hilly slopes, built-up zones, and veg-

etation-covered terrain, offering varied spatial contexts to assess geolocation accuracy under different surface and elevation conditions.

To ensure high-precision referencing and validation, GCPs were collected using a Trimble R12 Dual-Frequency DGPS receiver, operating in static mode. The base station was established on a known Benchmark (BM) located within the study area, providing a common geodetic reference for both the aerial data collected by the UAV and the ground survey points.

The choice of this study area is significant not only due to the terrain-induced challenges in placing GCPs but also because of the growing need for accurate geospatial datasets for watershed management, reservoir siltation monitoring, infrastructure planning, and disaster preparedness in the northeastern region of India [12]-[15]. The location map of the study area, along with GCP distributions and the UAV flight planning and geotagging using QBase 3D, is provided in Figure 1 and Figure 2, respectively.

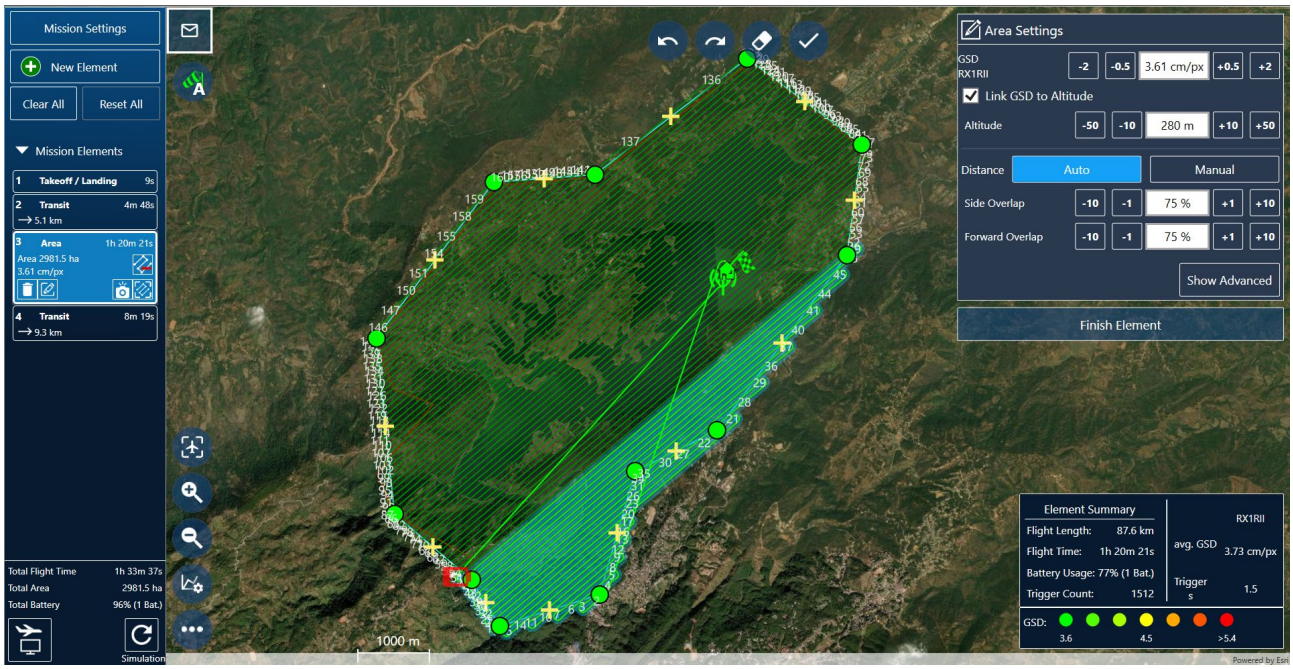


Figure 2. UAV flight planning and geotagging using QBase 3D.

3. Materials and Methods

This study involved the acquisition, processing, and evaluation of high-resolution geospatial datasets using a combination of UAV systems, GNSS survey equipment, and advanced photogrammetric software. The overall workflow consisted of two parallel data processing approaches: one based on PPK positioning and the other using GCPs acquired through Differential GPS (DGPS) survey. To ensure consistency and facilitate accurate comparison, all other processing parameters were kept constant between the two workflows, followed by machine learning-based error pattern analysis. The methodology adopted in this study is illustrated in Figure 3 below.

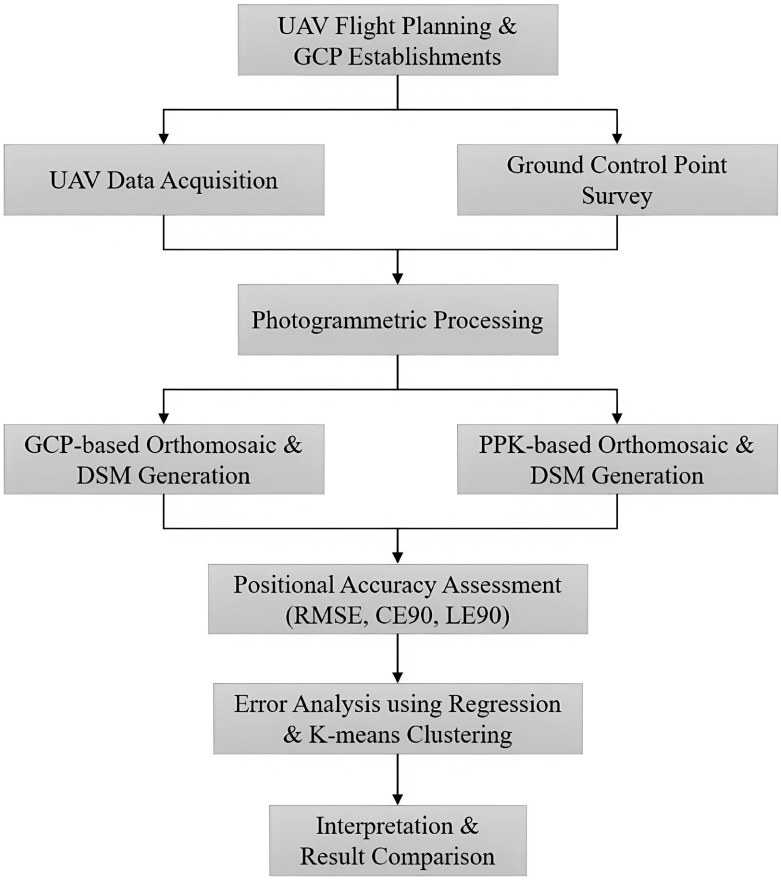


Figure 3. Workflow of the methodology adopted for UAV-based Photogrammetry, PPK correction, and machine learning-assisted accuracy assessment.

3.1. UAV Platform and Camera

The aerial surveys were conducted using the Trinity F90+ UAV, a VTOL fixed-wing platform developed by Quantum-Systems GmbH. The UAV offers long endurance (up to 90 minutes) and is well-suited for large-area mapping in complex terrain due to its autonomous navigation and terrain-following capabilities [7].

The onboard imaging system included a Sony RX1R II full-frame mirrorless camera with a 42.4 MP resolution and a fixed 35 mm lens. The camera's high-resolution sensor enabled detailed image capture with Ground Sampling Distances (GSDs) of less than 5 cm at typical flight altitudes of 120 - 150 meters AGL.

The UAV, equipped with a PPK-capable GNSS receiver, allows precise recording of camera positions during flight. Flight planning ensured at least 80% forward and 70% side overlap to support high-quality image matching during photogrammetric processing to ensure sufficient redundancy for robust SfM reconstruction.

3.2. Ground Control and Check Points

A total of ten GCPs were established across the study area to support georeferencing and accuracy assessment. These GCPs were surveyed using a Trimble R12

GNSS receiver operated in static mode (Figure 5), with reference to a known benchmark located within the site boundary. The GCPs were distributed across locations with varying vegetation conditions ranging from open to moderately vegetated areas, while fully dense canopy zones were avoided to maintain reliable DGPS observations. Each GCP was marked on the ground using clearly visible targets (Figure 4 & Figure 5) and their coordinates were recorded in UTM Zone 46N (WGS 84 datum). The GCPs served dual roles in this study: 1) as control points for orthomosaic generation in the GCP-based workflow, and 2) as reference points for positional accuracy assessment in both GCP and PPK-based outputs.

In accordance with the ASPRS Positional Accuracy Standards for Digital Geospatial Data [9], the GCPs were planned and surveyed to achieve horizontal and vertical accuracy within the 5 - 10 cm range, appropriate for high-resolution mapping. However, due to logistical constraints and time limitations, Independent Check Points (ICPs) could not be collected separately. As a result, the same GCPs were used for evaluating positional accuracy, which is acknowledged as a limitation of the study.

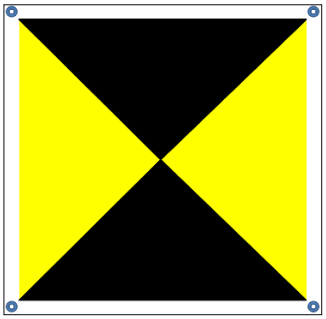


Figure 4. GCP ground marker used for field identification.

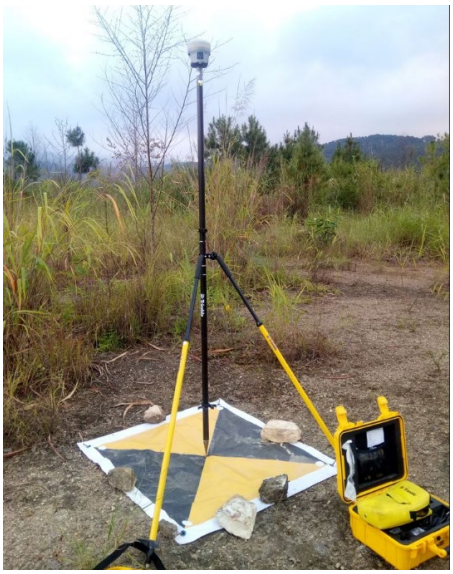


Figure 5. Trimble R12 GNSS receiver setup during static GCP survey.

3.3. Photogrammetric Processing and Software Tools

Two independent photogrammetric workflows were developed for generating georeferenced orthomosaics and DSMs for comparative evaluation:

- **GCP-based workflow:** Imagery was processed in Agisoft Metashape using all GCPs as control points for bundle adjustment and georeferencing.
- **PPK-based workflow:** Imagery was processed using onboard GNSS observations corrected through PPK techniques, without GCP input.

In both workflows, the camera calibration was allowed to adjust dynamically during bundle adjustment, and the same reprojection error thresholds were applied. The captured UAV images were processed in Agisoft Metashape Professional Edition (v2.0), following a consistent workflow for both the PPK-based and GCP-based cases. The key processing steps included:

Table 1. Summary of UAV specifications, ground survey setup, and processing tools.

Category	Parameter	Description
UAV Platform	UAV Model	Trinity F90+ (Quantum-Systems GmbH)
	UAV Type	Fixed-wing VTOL (Vertical Take-off and Landing)
	Endurance	Up to 90 minutes
	Positioning System	GNSS with PPK capability
	Flight Altitude	~120 meters AGL
	Flight Overlap	80% (forward), 70% (side)
	GNSS Base Station	Trimble R12 GNSS Receiver on known benchmark
Camera System	Sensor Model	Sony RX1R II
	Sensor Type	Full-frame CMOS
	Resolution	42.4 Megapixels
	Focal Length	35 mm fixed lens
	Image Format	JPEG/RAW
	Ground Sampling Distance (GSD)	~3 - 5 cm (depending on altitude)
Ground Survey	Survey Instrument	Trimble R12 GNSS Receiver (Dual Frequency)
	Operating Mode	Static mode with post-processing
	Ground Control Points (GCPs)	10 GCPs used in georeferencing
	Independent Check Points (ICPs)	Not used
	Coordinate System	WGS 84/UTM Zone 46N
Software Tools	Photogrammetry	Agisoft Metashape Professional v2.0
	GIS and Visualization	ArcGIS Pro 3.x, QGIS 3.32
	Accuracy Computation	Microsoft Excel (RMSE calculations)
	ML Analysis	Python (v3.10) with scikit-learn
Survey Scope	Area Covered	~25 square kilometers
	Terrain Type	Mixed terrain with slopes, vegetation, waterbody margins

- Image alignment and camera calibration (Structure-from-Motion);
- Dense point cloud generation;
- DSM and Digital Terrain Model (DTM) extraction;
- Orthomosaic generation;
- Georeferencing using either PPK camera positions or GCPs.

ArcGIS Pro 3.5 and QGIS 3.32 were used for spatial visualization and further accuracy assessment, including residual error computation and statistical summaries. Accuracy metrics were calculated using RMSE formulas in Excel and validated against ASPRS guidelines.

To explore potential ML-assisted insights, Python (v3.10) with scikit-learn was used for basic regression analysis and unsupervised clustering on spatial error datasets. These tools allowed analysis of relationships between terrain variables (e.g., slope, elevation) and geolocation errors [10]-[12]. **Table 1** provides a consolidated summary of the key equipment, technical specifications, and software tools employed in the study. It includes details on the UAV platform, onboard payload, and the ground control survey setup using the Trimble R12 GNSS receiver. Additionally, **Table 1** lists the photogrammetric and GIS software used for image processing, orthomosaic generation, and positional accuracy analysis.

3.4. Accuracy Assessment

Due to the unavailability of separate ICPs, the accuracy assessment was performed using the same 10 GCPs by comparing their known ground-truth coordinates with their corresponding photogrammetric locations in both datasets (PPK and GCP-based orthomosaics). The Root Mean Square Error (RMSE) was computed in the X (easting), Y (northing), and Z (elevation) directions using the following standard formula (Equation (1)):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - X_i'\right)^2} \tag{1}$$

where,

X_i' = extracted coordinate from orthomosaic;

X_i = ground-truth coordinate (from DGPS);

n = number of points (10 GCPs).

Separate RMSE values were calculated for:

- *Horizontal RMSE* ($RMSE_{XY}$): using both X and Y components;
- *Vertical RMSE* ($RMSE_Z$): using Z values.

In addition to RMSE, CE90 and LE90 statistics were computed to express the 90% confidence level positional accuracy, as per ASPRS (2015) [9] guidelines (Equations (2) & (3)):

- CE90 (Circular Error at 90%) estimates horizontal positional accuracy:

$$\text{CE90} = 1.6449 \times \sqrt{\text{RMSE}_X^2 + \text{RMSE}_Y^2} \tag{2}$$

- LE90 (Linear Error at 90%) estimates vertical accuracy:

$$\text{LE90} = 1.6449 \times \text{RMSE}_Z \tag{3}$$

These metrics provide a probabilistic threshold within which 90% of positional errors are expected to lie, thereby offering a more interpretable confidence level for operational planning and geospatial data quality assessment. Although using the same points for georeferencing and evaluation may underestimate the true error (as acknowledged by ASPRS, 2015), this approach still allows a comparative analysis between PPK and GCP-based workflows under identical conditions.

3.5. Machine Learning-Assisted Error Pattern Analysis

The complexity of error behavior in UAV Photogrammetry, particularly in hilly and forested landscapes, often exceeds the explanatory power of conventional descriptive statistics. To address this limitation, we incorporated ML techniques as an exploratory step for modeling and interpreting positional error patterns in the PPK-based orthomosaic outputs. The goal was not only to quantify deviations but also to uncover systematic influences of terrain factors and to group GCPs with similar error characteristics.

Two categories of inputs were considered:

1) Error components: ΔX, ΔY, ΔZ derived by comparing UAV-based coordinates with DGPS ground truth.

2) Terrain variables: slope (S, in degrees), elevation (E), and aspect (A), extracted from the GCP-based DSM.

All variables were standardized to zero mean and unit variance before modeling to eliminate dimensional bias.

3.5.1. Regression Modeling

A linear regression model was applied to examine the influence of terrain slope (S) on vertical error (ΔZ) in the PPK-based workflow:

$$\Delta Z_{ppk,i} = \beta_0 + \beta_1 S_i + \epsilon_i \tag{4}$$

where:

- $\Delta Z_{ppk,i}$ = vertical error at GCP I;
- S_i = terrain slope at GCP I;
- β_0 = intercept;
- β_1 = slope coefficient;
- ϵ_i = residual error.

Model performance was quantified using the coefficient of determination:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(\Delta Z_{ppk,i} - \Delta \hat{Z}_{ppk,i} \right)^2}{\sum_{i=1}^{n} \left(\Delta Z_{ppk,i} - \Delta \overline{Z}_{ppk} \right)^2} \tag{5}$$

where $\Delta \hat{Z}_{ppk,i}$ is the predicted vertical error and $\Delta \overline{Z}_{ppk}$ the mean vertical error.

Given the limited dataset of ten GCPs, linear regression was chosen intentionally for its stability, robustness, and interpretability. More complex non-linear models were avoided to prevent overfitting and to ensure transparency in explaining how slope influences vertical error.

3.5.2. Clustering of Positional Errors

To identify groups of similar error behavior, K-means clustering was applied to three-dimensional error vectors:

$$e_i = [\Delta X_i, \Delta Y_i, \Delta Z_i] \tag{6}$$

The K-means objective is to partition n error vectors into k clusters by minimizing the within-cluster variance:

$$\min \sum_{j=1}^{k} \sum_{i \in C_j} \left\| e_i - \mu_j \right\|^2 \tag{7}$$

where:

- C_j = cluster j;
- μ_j = centroid of cluster j;
- $\|\cdot\|$ = Euclidean norm.

K-means was selected due to its computational simplicity, intuitive centroid-based grouping, and suitability for small datasets. This approach enables systematic grouping of GCPs into distinct error regimes, facilitating interpretation beyond global accuracy statistics.

4. Results

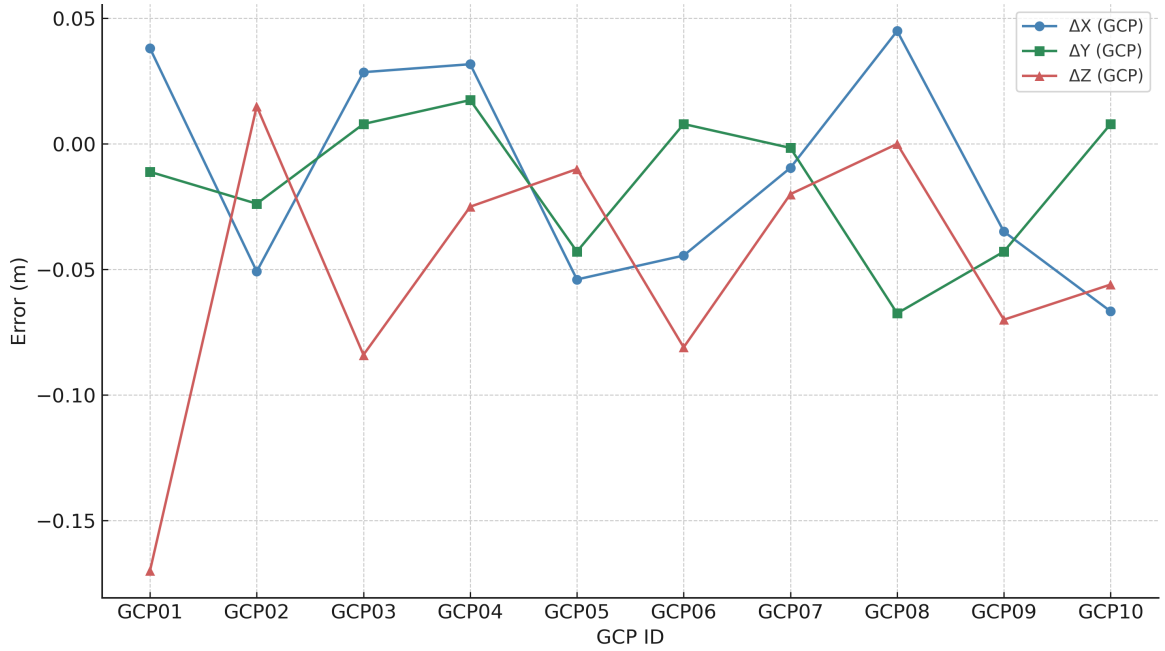
This section presents the comparative accuracy assessment between the GCP-based and PPK-based UAV workflows, followed by the machine learning-assisted analysis of error patterns. Results are structured into three subsections: 1) component-wise positional accuracy evaluation, 2) regression analysis of terrain slope versus PPK vertical error, and 3) clustering analysis of error behavior.

4.1. Positional Accuracy Comparison

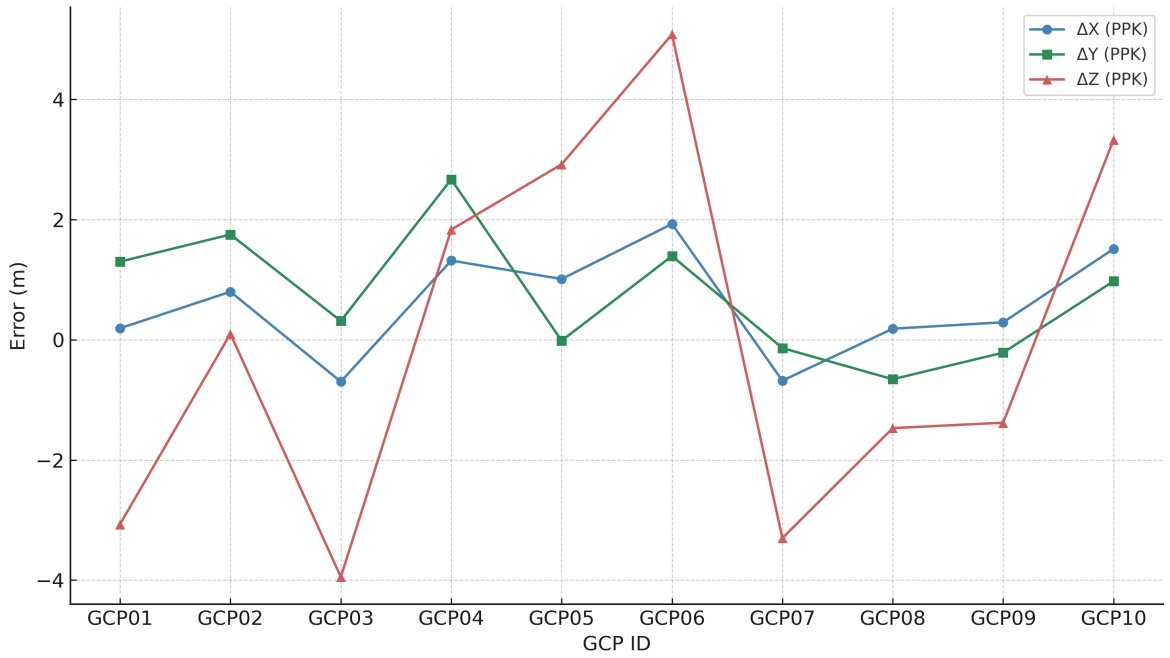
The positional accuracy assessment was carried out by comparing UAV-derived orthomosaic coordinates (from both GCP-based and PPK-based workflows) with DGPS-surveyed ground truth at 10 GCPs. Component-wise deviations in easting (ΔX), northing (ΔY), and elevation (ΔZ) were computed for each GCP (Table 2).

Table 2. Component-wise positional differences for each GCP in GCP-based and PPK-based orthomosaics.

GCP_ID	ΔX_GCP (m)	ΔY_GCP (m)	ΔZ_GCP (m)	ΔX_PPK (m)	ΔY_PPK (m)	ΔZ_PPK (m)
GCP01	0.038	−0.011	−0.17	0.192	1.301	−3.07
GCP02	−0.051	−0.024	0.015	0.8	1.752	0.1
GCP03	0.029	0.008	−0.084	−0.697	0.314	−3.948
GCP04	0.032	0.017	−0.025	1.319	2.669	1.833
GCP05	−0.054	−0.043	−0.01	1.011	−0.016	2.92
GCP06	−0.044	0.008	−0.081	1.927	1.394	5.084
GCP07	−0.01	−0.002	−0.02	−0.681	−0.138	−3.3
GCP08	0.045	−0.067	0	0.185	−0.657	−1.47
GCP09	−0.035	−0.043	−0.07	0.291	−0.214	−1.38
GCP10	−0.067	0.008	−0.056	1.512	0.975	3.33



(a)



(b)

Figure 6. (a) Component-wise positional errors (ΔX, ΔY, ΔZ) for each GCP based on the GCP-based orthomosaic. The results confirm sub-decimeter deviations across all GCPs, with minimal variation in planimetric and vertical components. (b) Component-wise positional errors (ΔX, ΔY, ΔZ) for each GCP based on the PPK-based orthomosaic. Substantially larger deviations are observed, especially in the vertical component (ΔZ), where errors exceed ±3 m for several GCPs located in steep or vegetated terrain.

For the GCP-based workflow, deviations were minimal across all dimensions. The RMSE values were 0.043 m (X), 0.031 m (Y), and 0.072 m (Z), with a combined horizontal $RMSE_{XY}$ of 0.053 m. The CE90 and LE90 values were 0.080 m

and 0.119 m, respectively, confirming sub-decimeter positional accuracy [9]. The line graph of component-wise deviations (**Figure 6(a)**) further illustrates the consistently low error levels across all GCPs. These results validate the reliability of conventional GCP-based UAV photogrammetry for precision mapping tasks.

In contrast, the PPK-based workflow exhibited significantly larger positional errors. RMSE values were 1.025 m (X), 1.240 m (Y), and 2.980 m (Z), with $RMSE_{XY}$ = 1.609 m. CE90 and LE90 values were 2.441 m and 4.902 m, respectively. Vertical errors were particularly pronounced, with several GCPs (e.g., GCP03, GCP06) exceeding ±3 m (**Figure 6(b)**). A direct comparison (**Figure 7**) highlights the sharp contrast in accuracy between the two workflows, with PPK-based results diverging significantly in both horizontal and vertical components.

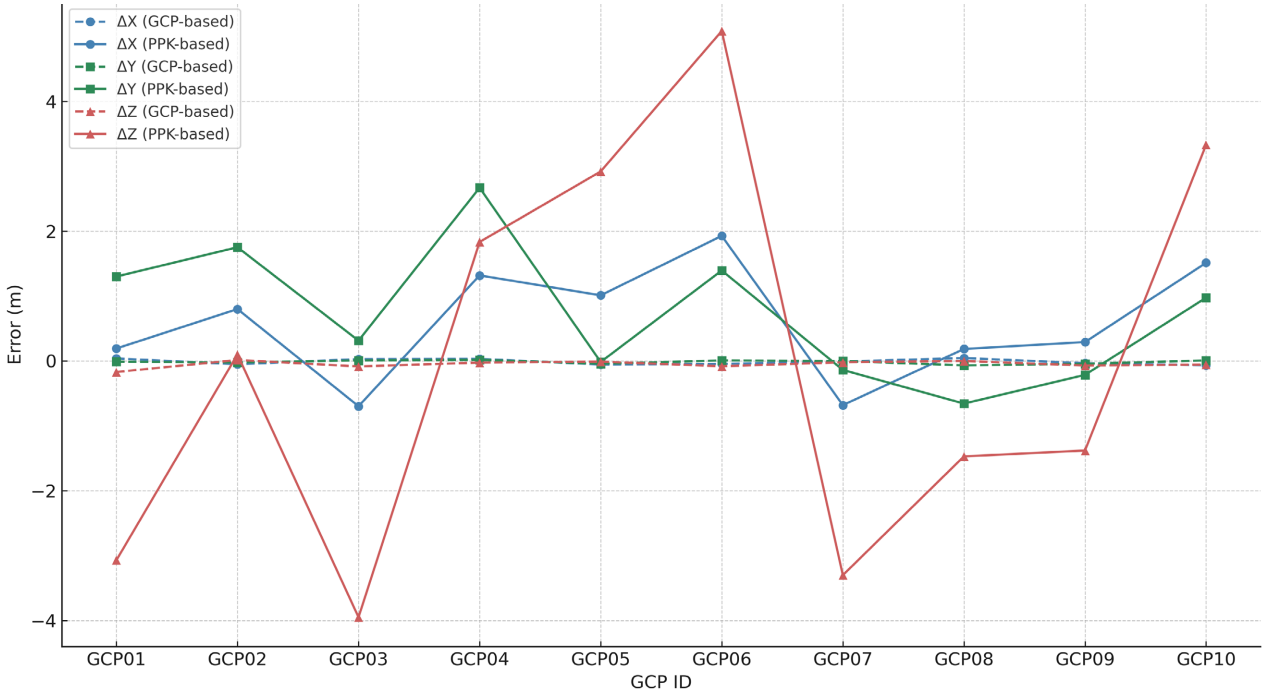


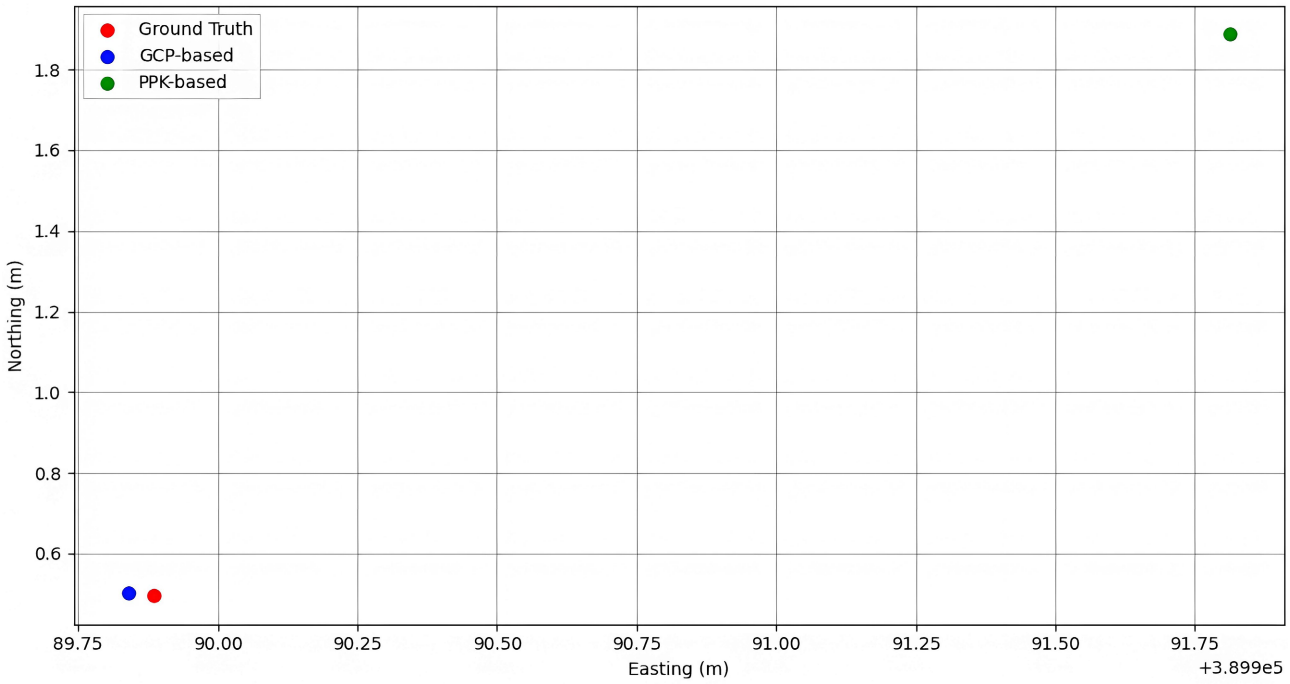
Figure 7. Comparison of component-wise positional errors (ΔX, ΔY, ΔZ) between GCP-based and PPK-based orthomosaics. While the GCP-based workflow consistently achieves near-zero deviations, the PPK-based workflow exhibits significant vertical inaccuracies and greater variability in planimetric components, underscoring the performance gap between the two approaches in complex terrain.

Table 3 summarizes the RMSE, CE90, and LE90 values for both workflows. **Figure 8(a)** & **Figure 8(b)** and **Figure 9(a)** & **Figure 9(b)** show the spatial distribution and the spatial offset visualization for GCP06 & GCP07, highlighting differences in horizontal placement between the two methods.

Table 3. Summary of RMSE, CE90, and LE90 for GCP-based and PPK-based orthomosaics.

Workflow	$RMSE_X$ (m)	$RMSE_Y$ (m)	$RMSE_Z$ (m)	$RMSE_{XY}$ (m)	CE90 (m)	LE90 (m)
GCP-based	0.043	0.031	0.072	0.053	0.080	0.119
PPK-based	1.025	1.240	2.980	1.609	2.441	4.902

(a)

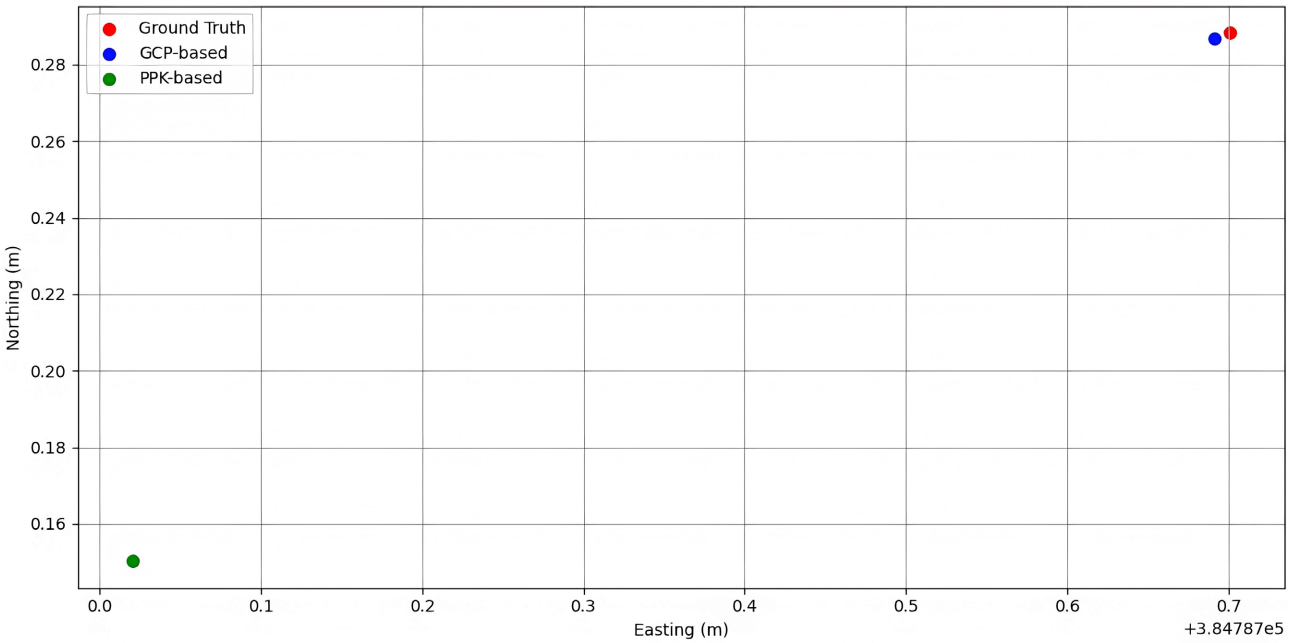


(b)

Figure 8. (a) Zoomed-in spatial distribution of GCP06 overlaid on the GCP-based orthomosaic image. The red cross indicates the DGPS-surveyed position, blue circle represents the GCP-based orthomosaic output, and green circle shows the PPK-based orthomosaic position. (b) Spatial offset visualization for GCP06.

These results highlight a trade-off: while PPK reduces the need for extensive GCP deployment, it introduces positional inaccuracies, particularly in elevation. This is especially critical in hilly terrain, where relief displacement, GNSS signal obstruction, and image matching inconsistencies are more pronounced [4] [8].

(a)

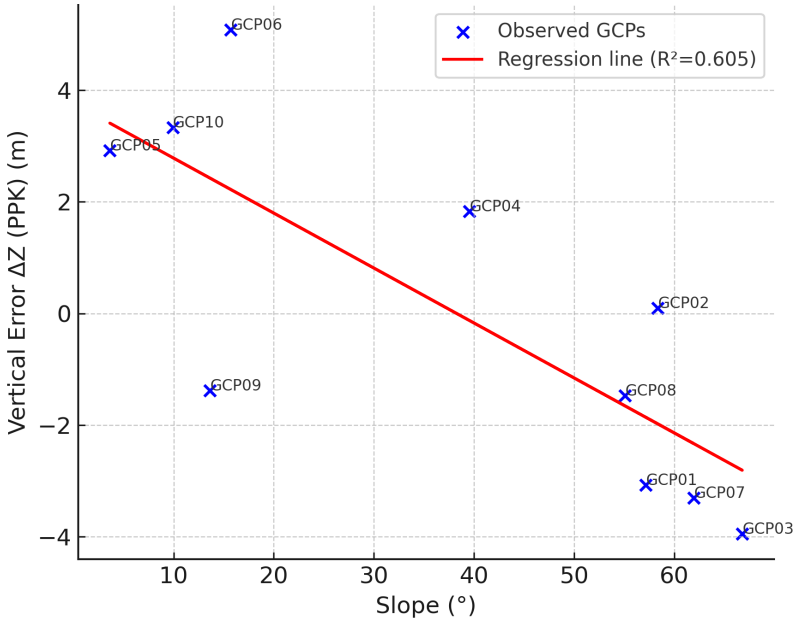


(b)

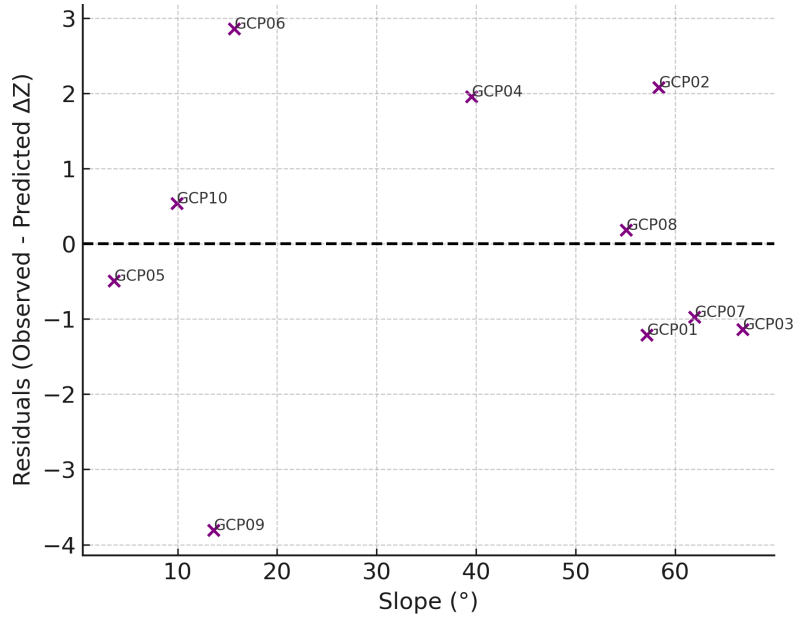
Figure 9. (a) Zoomed-in spatial distribution of GCP07 overlaid on the GCP-based orthomosaic image. (b) Spatial offset visualization for GCP07 with noticeable positional deviation in the PPK-based result.

4.2. Regression Analysis: Terrain Slope vs. PPK Vertical Error

A linear regression model (Equation (4)) was used to examine the relationship between terrain slope (˚) and vertical error (ΔZ) from the PPK-based workflow. The fitted regression line (**Figure 10(a)**) yielded a coefficient of determination of $R^2 = 0.605$, indicating that slope explains approximately 60% of the variance in vertical error.



(a)



(b)

Figure 10. (a) Regression plot between terrain slope (˚) and vertical error (ΔZ) from the PPK-based workflow. The fitted regression model (Equation (4)) produced a coefficient of determination of R^2 = 0.605, indicating that slope explains about 60% of the variance in vertical error. Annotated points denote individual GCPs. (b) Residual plot of the regression model between slope and PPK-derived vertical error (ΔZ). The residuals display moderate scatter, with a discernible trend, confirming that slope is an important, though not exclusive, determinant of vertical error. Other contributing factors include vegetation density and GNSS multipath effects.

This result suggests a moderately strong relationship, with larger slopes generally associated with greater vertical deviations. However, variability among individ-

ual GCPs implies that other factors, such as vegetation cover, GNSS multipath, and tie-point distribution, also influence vertical error.

Residual analysis (Figure 10(b)) showed moderate scatter, reinforcing that slope is a significant but not exclusive driver of PPK error.

4.3. Clustering Analysis of PPK Error Patterns

K-means clustering was applied to the three-dimensional PPK error vectors (ΔX, ΔY, ΔZ) to identify systematic patterns of deviation. The analysis revealed three distinct clusters (Figure 11), with centroids presented in Table 4.

Table 4. Cluster centroids of PPK-based positional errors.

Cluster	ΔX (m)	ΔY (m)	ΔZ (m)
Cluster 0	−0.142	0.121	−2.634
Cluster 1	1.484	0.785	3.778
Cluster 2	1.059	2.210	0.967

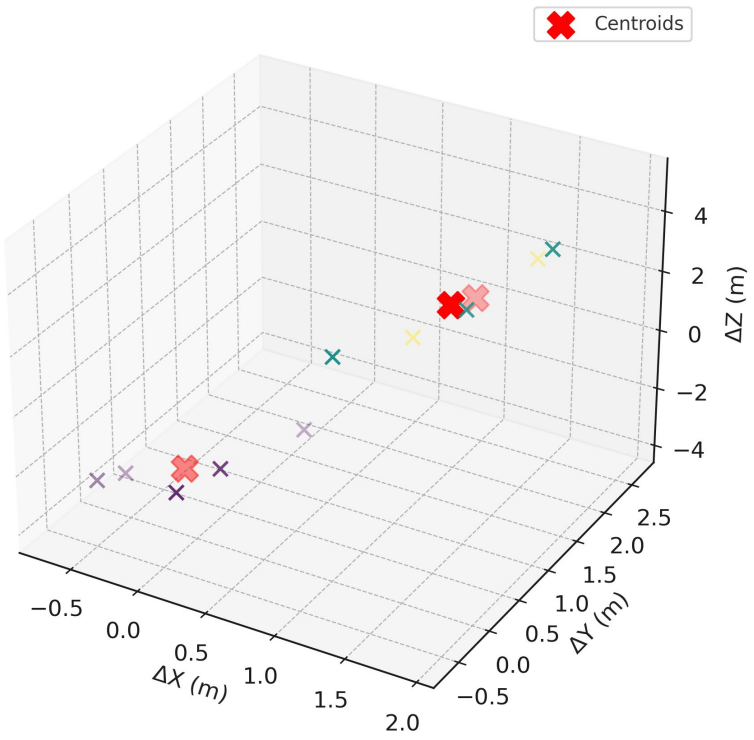


Figure 11. Three-dimensional K-means clustering of PPK positional errors using ΔX, ΔY, and ΔZ components. GCPs are color-coded according to cluster assignment, and red "×" markers indicate cluster centroids. Cluster 0 is dominated by systematic vertical underestimation ($\Delta Z \approx -2.63$ m), Cluster 1 by vertical overestimation ($\Delta Z \approx +3.78$ m), and Cluster 2 by horizontal offsets ($\Delta X \approx +1.06$ m, $\Delta Y \approx +2.21$ m), highlighting distinct error regimes in the PPK workflow.

- *Cluster 0*: Systematic vertical underestimation ($\Delta Z \approx -2.63$ m) with negligible horizontal offsets.

- *Cluster 1*: Systematic vertical overestimation ($\Delta Z \approx +3.78$ m), representing the largest vertical errors.
- *Cluster 2*: Significant horizontal offsets ($\Delta X \approx +1.06$ m, $\Delta Y \approx +2.21$ m) with moderate vertical error ($\Delta Z \approx +0.97$ m).

The clustering results highlight the heterogeneous nature of PPK errors in complex terrain. While some GCPs experienced systematic vertical bias, others exhibited horizontal drift, suggesting multiple interacting sources of error. Unlike global RMSE values, clustering enables a more nuanced understanding of positional error regimes.

This analysis highlights the presence of systematic error trends in the PPK output and supports the hypothesis that such patterns may arise from environmental, geometric, or sensor-specific influences during data acquisition.

5. Discussion

The findings of this study provide valuable insights into the performance of GCP-based and PPK-based UAV Photogrammetry in complex terrain and demonstrate the potential of machine learning approaches for interpreting positional error patterns. The GCP-based workflow consistently achieved sub-decimetre accuracy in all dimensions, reaffirming its reliability for high-precision applications such as cadastral mapping, infrastructure planning, and detailed topographic modelling. These results align with earlier studies demonstrating the robustness of GCP-based georeferencing when an adequate number of well-distributed GCPs are deployed [16]. Nonetheless, the need for field-intensive GCP surveys remains a practical limitation in steep, vegetated, or otherwise inaccessible environments such as those surrounding the Umiam Reservoir.

In contrast, the PPK-based workflow reduced the requirement for extensive ground control but exhibited noticeably higher positional errors, particularly in the vertical dimension. With $RMSE_Z$ values approaching 3 m and LE90 estimates exceeding 4.9 m, the vertical accuracy of PPK-only outputs is insufficient for elevation-sensitive applications such as hydrological modelling, reservoir management, or slope stability analysis. These observations are consistent with prior research indicating that although PPK can achieve accuracy comparable to GCP-based workflows under favourable conditions [3] [17], its performance degrades in rugged or forested terrain where relief displacement, GNSS signal obstruction, and multipath interference are more prevalent.

The regression analysis adds further clarity by quantifying the influence of terrain slope on PPK-derived vertical error. With slope accounting for approximately 60% of the variance in ΔZ, the results confirm that steeper terrain amplifies vertical inaccuracies. Importantly, slope often interacts with other terrain-related factors, such as vegetation density, surface roughness, and reduced sky visibility, that collectively intensify multipath effects and degrade GNSS corrections. These interactions help explain the terrain-error relationship observed in this study. The implication is that slope and its associated terrain conditions should be explicitly

considered when planning UAV surveys in hilly regions, as vertical errors can be anticipated based on local morphology.

The clustering analysis further enriches the interpretation by revealing distinct error regimes: 1) systematic vertical underestimation, 2) systematic vertical overestimation, and 3) horizontal drift. This demonstrates that positional errors are not uniformly distributed but follow identifiable patterns linked to local conditions. Such findings extend beyond conventional RMSE-based evaluations by providing a more nuanced understanding of error heterogeneity. This methodological contribution, using unsupervised machine learning to characterize positional error regimes, represents a novel aspect of this study. It shows how relatively simple ML models can reveal systematic biases that are otherwise obscured in global accuracy statistics.

From an operational perspective, the results highlight both the benefits and limitations of PPK workflows in complex terrain. On the one hand, PPK enables efficient data acquisition without requiring dense GCP networks, which is particularly advantageous in inaccessible or hazardous environments. On the other hand, the observed error magnitudes indicate that PPK alone cannot ensure mapping-grade accuracy in hilly regions. A promising pathway is the adoption of hybrid workflows, where sparse GCPs are strategically placed in areas prone to error (e.g., steep slopes or dense vegetation) to constrain PPK outputs, or where machine learning models are trained to predict and correct terrain-induced errors.

Finally, these findings hold broader implications for geospatial practices in the North Eastern Region of India and similar hilly environments worldwide. For critical applications such as reservoir monitoring, landslide risk assessment, or infrastructure development, accuracy requirements may exceed what PPK workflows can provide on their own. By integrating PPK with minimal ground control and ML-assisted diagnostics, it is possible to achieve a balance between operational efficiency and mapping reliability.

6. Conclusions

This study assessed the positional accuracy of UAV-derived orthomosaics using GCP-based and PPK-based workflows in the complex terrain of the Umiam Reservoir, Meghalaya, and applied machine learning to analyze error patterns. The results demonstrate that the GCP-based workflow achieved sub-decimeter accuracy ($RMSE_Z$ = 0.072 m, LE90 = 0.119 m), confirming its reliability for precision mapping. In contrast, the PPK-based workflow introduced substantial deviations ($RMSE_Z$ = 2.98 m, LE90 = 4.90 m), particularly in the vertical dimension, limiting its suitability for applications requiring fine elevation accuracy.

The machine learning analyses provided additional insights beyond conventional accuracy metrics. Regression analysis showed that slope explained approximately 60% of the variance in vertical error, highlighting terrain steepness as a key driver of PPK inaccuracies. Clustering analysis identified three distinct error regimes: systematic vertical underestimation, vertical overestimation, and hori-

zontal drift, demonstrating that positional errors are heterogeneous and context dependent. These findings underscore the value of ML-assisted diagnostics in identifying systematic error behavior that conventional RMSE values cannot capture.

At the same time, certain limitations of this study must be acknowledged. The analysis was based on 10 GCPs, which limits the statistical robustness of the ML models. The focus on a single case study area restricts the generalizability of results to other terrains. Moreover, only slope was explicitly modeled as a terrain predictor, while other variables such as aspect, vegetation cover, or GNSS satellite geometry were not considered. Finally, the study employed basic ML models (linear regression, K-means clustering), which are effective for exploratory analysis but less powerful than advanced non-linear approaches.

These limitations, however, define clear directions for future research. Expanding the dataset to include more GCPs and multiple study areas will enhance model robustness and generalizability. Incorporating additional predictors such as vegetation indices, canopy height, or GNSS satellite visibility will provide a more comprehensive understanding of error sources. Advanced machine learning approaches, such as Random Forests, Gradient Boosting, or Neural Networks, should be tested to capture complex, non-linear relationships. Operationally, hybrid workflows that combine PPK with sparse GCP deployment or terrain-aware ML corrections appear promising for balancing efficiency and accuracy.

While PPK-based workflows offer operational efficiency by reducing dependence on GCPs, their positional accuracy is compromised in rugged terrain. By integrating PPK with limited ground control and ML-based error diagnostics, UAV Photogrammetry can be made both efficient and reliable. This study contributes methodological innovation and practical recommendations, with broader implications for UAV mapping in hilly regions of Northeast India and other complex terrains worldwide.

Acknowledgements

The authors acknowledge the support provided by the North Eastern Space Applications Centre (NESAC), Department of Space, Government of India, for facilitating UAV data acquisition, DGPS survey, and photogrammetric processing using licensed software. We also extend our appreciation to the Meghalaya Forest Department for granting access and providing logistical support during field operations around the Umiam Reservoir. Special thanks to the technical teams involved in UAV operation and ground control surveys, whose efforts were critical to the success of this study. The insights gained from this work contribute to ongoing efforts in enhancing geospatial accuracy for UAV applications in complex terrain settings.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Colomina, I. and Molina, P. (2014) Unmanned Aerial Systems for Photogrammetry and Remote Sensing: A Review. *ISPRS Journal of Photogrammetry and Remote Sensing*, **92**, 79-97. https://doi.org/10.1016/j.isprsjprs.2014.02.013

[2] Verhoeven, G. (2011) Taking Computer Vision Aloft-Archaeological Three-Dimensional Reconstructions from Aerial Photographs with Photoscan. *Archaeological Prospection*, **18**, 67-73. https://doi.org/10.1002/arp.399

[3] Martínez-Carricondo, P., Agüera-Vega, F. and Carvajal-Ramírez, F. (2023) Accuracy Assessment of RTK/PPK UAV-Photogrammetry Projects Using Differential Corrections from Multiple GNSS Fixed Base Stations. *Geocarto International*, **38**, 1-21.

[4] Zhao, H., Li, G., Chen, Z., Zhang, S., Zhang, B. and Cheng, X. (2024) Impacts of GCP Distributions on UAV-PPK Photogrammetry at Sermeq Avannarleq Glacier, Greenland. *Remote Sensing*, **16**, Article 3934. https://doi.org/10.3390/rs16213934

[5] Mesas-Carrascosa, F.J., Torres-Sánchez, J., Clavero-Rumbao, I., García-Ferrer, A., Peña, J.M., Borra-Serrano, I., *et al.* (2015) Assessing Optimal Flight Parameters for Generating Accurate Multispectral Orthomosaicks by UAV to Support Site-Specific Crop Management. *Remote Sensing*, **7**, 12793-12814. https://doi.org/10.3390/rs71012793

[6] Turner, D., Lucieer, A. and Watson, C. (2012) An Automated Technique for Generating Georectified Mosaics from Ultra-High Resolution Unmanned Aerial Vehicle (UAV) Imagery, Based on Structure from Motion (SfM) Point Clouds. *Remote Sensing*, **4**, 1392-1410. https://doi.org/10.3390/rs4051392

[7] Quantum-Systems (2023) Trinity F90+ Product Specifications. https://quantum-systems.com/

[8] Agüera-Vega, F., Carvajal-Ramírez, F. and Martínez-Carricondo, P. (2017) Assessment of Photogrammetric Mapping Accuracy Based on Variation Ground Control Points Number Using Unmanned Aerial Vehicle. *Measurement*, **98**, 221-227. https://doi.org/10.1016/j.measurement.2016.12.002

[9] ASPRS (2015) ASPRS Positional Accuracy Standards for Digital Geospatial Data. American Society for Photogrammetry and Remote Sensing. https://www.asprs.org/

[10] Yang, J., Xu, J., Lv, Y., Zhou, C., Zhu, Y. and Cheng, W. (2023) Deep Learning-Based Automated Terrain Classification Using High-Resolution DEM Data. *International Journal of Applied Earth Observation and Geoinformation*, **118**, Article ID: 103249. https://doi.org/10.1016/j.jag.2023.103249

[11] Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G. and Johnson, B.A. (2019) Deep Learning in Remote Sensing Applications: A Meta-Analysis and Review. *ISPRS Journal of Photogrammetry and Remote Sensing*, **152**, 166-177. https://doi.org/10.1016/j.isprsjprs.2019.04.015

[12] Uhl, J.H. and Leyk, S. (2023) Spatially Explicit Accuracy Assessment of Deep Learning-Based, Fine-Resolution Built-Up Land Data in the United States. *International Journal of Applied Earth Observation and Geoinformation*, **123**, Article ID: 103469. https://doi.org/10.1016/j.jag.2023.103469

[13] Seo, D.M., Woo, H.J., Hong, W.H., Seo, H. and Na, W.J. (2024) Optimization of Number of GCPs and Placement Strategy for UAV-Based Orthophoto Production. *Applied Sciences*, **14**, Article 3163. https://doi.org/10.3390/app14083163

[14] Saikhom, V. and Kalita, M. (2024) UAV for Remote Sensing Applications: An Analytical Review. In: *Lecture Notes in Electrical Engineering*, Springer, 51-59. https://doi.org/10.1007/978-981-99-4362-3_6

[15] JOUAV (2024) A Comprehensive Guide to Using Ground Control Points for Drone Surveying. https://www.jouav.com

[16] Zhong, H., Duan, Y., Tao, P. and Zhang, Z. (2025) Influence of Ground Control Point Reliability and Distribution on UAV Photogrammetric 3D Mapping Accuracy. *Geo-Spatial Information Science*, **28**, 1998-2018. https://doi.org/10.1080/10095020.2025.2451204

[17] Lee, T., Bettinger, P., Merry, K. and Cieszewski, C. (2023) The Effects of Nearby Trees on the Positional Accuracy of GNSS Receivers in a Forest Environment. *PLOS ONE*, **18**, e0283090. https://doi.org/10.1371/journal.pone.0283090